

# The Large Deviation Principle in Statistical Mechanics: an Expository Account

J.T. Lewis  
Dublin Institute for Advanced Studies  
10 Burlington Road  
Dublin 4, Ireland

## §1 Introduction

At the 1983 Swansea meeting I described joint work with J.V. Pulè [1] on the weak law of large numbers in statistical mechanics; at that time we were unaware of its connection with the principle of large deviations as formulated by Varadhan [2]. Since then, we have found Varadhan's theorems to be a powerful tool in statistical mechanics (results of use in investigating models of an interacting boson gas are described elsewhere in this volume [3]) and so it may prove worthwhile to look again at the results of [1] with the benefit of hindsight.

## §2 The Grand Canonical Pressure

Since the pioneering work of van Hove [4], the importance of proving the existence of thermodynamic functions in the thermodynamic limit has been recognized. We recall the definition of the grand canonical pressure: consider a sequence  $\{\Lambda_\ell : \ell = 1, 2, \dots\}$  of regions of Euclidean space  $\mathbb{R}^d$ , and denote the volume of  $\Lambda_\ell$  by  $V_\ell$ ; associated with each region  $\Lambda_\ell$  is a countable set  $\Omega_\ell$ , the set of configurations of particles in  $\Lambda_\ell$ ; on  $\Omega_\ell$  are defined random variables  $H_\ell : \Omega_\ell \rightarrow \mathbb{R}$  and  $N_\ell : \Omega_\ell \rightarrow \mathbb{N}$ ;  $H_\ell(\omega)$  is interpreted as the energy of the configuration  $\omega$  and  $N_\ell(\omega)$  as the number of particles in  $\omega$ . The grand canonical measure  $P_\ell^\mu$  with chemical potential  $\mu$  is defined on subsets of  $\Omega_\ell$  by

$$P_\ell^\mu[A] = \Xi_\ell(\mu)^{-1} \sum_{\omega \in A} e^{\beta\{\mu N_\ell(\omega) - H_\ell(\omega)\}}, \quad (2.1)$$

here  $\beta = 1/kT$  is the inverse temperature and  $\Xi_\ell(\mu)$  is the grand canonical partition function given by

$$\Xi_\ell(\mu) = \sum_{\omega \in \Omega_\ell} e^{\beta\{\mu N_\ell(\omega) - H_\ell(\omega)\}}. \quad (2.2)$$

The grand canonical pressure  $p_\ell(\mu)$  is defined by

$$p_\ell(\mu) = (\beta V_\ell)^{-1} \ln \Xi_\ell(\mu). \quad (2.3)$$

It is closely related to the cumulant generating function for the particle number density  $X_\ell = N_\ell/V_\ell$ ; a straightforward calculation yields the formula

$$\int_{[0,\infty)} e^{\beta V_\ell x t} K_\ell^\mu[dx] = e^{\beta V_\ell \{p_\ell(\mu+t) - p_\ell(\mu)\}}, \quad (2.4)$$

where  $K_\ell^\mu$  is the distribution function for  $X_\ell$  defined by  $K_\ell^\mu = P_\ell^\mu \circ X_\ell^{-1}$ . We have introduced the concepts associated with the grand canonical pressure in the simplest case, namely, when  $\Omega_\ell$  is a countable set. But the formula (2.4), linking the distribution function  $K_\ell^\mu$  with the grand canonical pressure holds in wider contexts; for example, when  $\Omega_\ell$  is an arbitrary measure space carrying a pair of random variables  $H_\ell$  and  $N_\ell$  from which a grand canonical Gibbs measure can be defined, or when  $H_\ell$  and  $N_\ell$  are commuting self-adjoint operators on some hilbert space  $\mathcal{H}_\ell$  such that trace  $(e^{\beta\{\mu N_\ell - H_\ell\}})$  is finite. Our first assumption in the remainder of this note is that the distribution function  $K_\ell^\mu$  satisfies (2.4) for some function  $p_\ell(\mu)$ . Our second assumption is that  $p(\mu) = \lim_{\ell \uparrow \infty} p_\ell(\mu)$  exists. From these two assumptions, much follows: the large deviation upper bound holds and, with it, the Berezin-Sinai criterion for a first-order phase-transition; the large deviation lower bound holds in the complement of first-order phase-transition segments; if the limit function  $\mu \mapsto p(\mu)$  is differentiable at some  $\mu_0$  then  $\{K_\ell^{\mu_0} : \ell = 1, 2, \dots\}$  converges in distribution to the degenerate distribution concentrated at  $p'(\mu_0)$ .

In probability theory, it is natural to prove large deviation results for sums of independent, or weakly dependent, random variables: the Markov chain condition is an example of a condition of weak dependence. I claim that, in statistical mechanics, the natural condition of weak dependence is the existence of the pressure. In this note, we explore the consequences for the distribution of the particle number density of the existence of the pressure.

### §3 The General Setting

For each  $\mu$  in some open interval  $D$  of the real line, let  $\{K_\ell^\mu : \ell = 1, 2, \dots\}$  be a sequence of probability measures on  $[0, \infty)$  satisfying

(P1)

$$\int_{[0,\infty)} e^{V_\ell t x} K_\ell^\mu[dx] = e^{V_\ell \{p_\ell(\mu+t) - p_\ell(\mu)\}} < \infty,$$

where  $\{V_\ell : \ell = 1, 2, \dots\}$  is a sequence of positive constants diverging to  $+\infty$ .

(P2) The limit  $p(\mu) = \lim_{\ell \uparrow \infty} p_\ell(\mu)$  exists for all values of  $\mu$  in the interval of definition,  $D$ .

The first consequence makes use of Hölder's Inequality and we omit the proof:

Lemma 1

Assume that (P1) holds; then  $\mu \mapsto p_\ell(\mu)$  is convex. Assume, in addition, that (P2) holds; then  $\mu \mapsto p(\mu)$  is convex.

Lemma 2

For all  $\mu$  and  $\mu + \alpha$  in the domain of definition of  $K_\ell$ , the measures  $K_\ell^\mu$  and  $K_\ell^{\mu+\alpha}$  are mutually absolutely continuous:

$$K_\ell^{\mu+\alpha}[dx] = e^{V_\ell c_\ell^\mu(x; \alpha)} K_\ell^\mu[dx], \quad (3.1)$$

where

$$c_\ell^\mu(x; \alpha) = \alpha x + p_\ell(\mu) - p_\ell(\mu + \alpha). \quad (3.2)$$

Proof:

$$\int_{[0, \infty)} e^{xt} \cdot e^{V_\ell \alpha x} K_\ell^\mu[dx] = e^{V_\ell \{p_\ell(\mu + \alpha + \frac{t}{V_\ell}) - p_\ell(\mu)\}},$$

by (P1); again, by (P1), we have

$$\begin{aligned} & \frac{e^{V_\ell p_\ell(\mu + \alpha)}}{e^{V_\ell p_\ell(\mu)}} \int_{[0, \infty)} e^{xt} K_\ell^{\mu+\alpha}[dx] \\ &= e^{V_\ell \{p_\ell(\mu + \alpha + \frac{t}{V_\ell}) - p_\ell(\mu)\}}. \end{aligned}$$

The claim follows from the uniqueness theorem for Laplace transforms.

Theorem 1

Assume that (P1) and (P2) hold and that  $p$  is differentiable at  $\mu$ ; then  
 (1) the limit  $\rho = \lim_{\ell \uparrow \infty} \int_{[0, \infty)} x K_\ell^\mu[dx]$  exists.  
 (2) the sequence  $\{K_\ell^\mu : \ell = 1, 2, \dots\}$  converges weakly to the degenerate distribution  $\delta_\rho$  concentrated at  $\rho$ .

The proof utilises the convexity of the functions  $p_\ell$ ,  $\ell = 1, 2, \dots$ , established in Lemma 1. This enables us to apply

### Griffith's Lemma

Let  $\{f_\ell : \ell = 1, 2, \dots\}$  be a sequence of convex functions defined on a common open interval  $G$  converging pointwise to a function  $f$ . Let  $\{x_\ell : \ell = 1, 2, \dots\}$  be a sequence of points of  $G$  converging to a point  $x$  of  $G$ . Then

$$f'_-(x) \leq \liminf_{\ell \rightarrow \infty} (f_\ell)'_-(x_\ell) \leq \limsup_{\ell \rightarrow \infty} (f_\ell)'_+(x_\ell) \leq f'_+(x).$$

(See [1] and references contained therein.)

### Proof of Theorem 1:

By an elementary computation,  $\int_{[0, \infty)} x K_\ell^\mu[dx] = p'_\ell(\mu)$  since, by (P1), the moment generating function  $s \mapsto \int_{[0, \infty)} e^{sx} K_\ell^\mu[dx]$  of  $K_\ell^\mu$  is finite on a neighbourhood of zero. Since  $p$  is assumed to be differentiable at  $\mu$ , it follows from Griffith's Lemma that  $\{p'_\ell(\mu) : \ell = 1, 2, \dots\}$  converges to  $p'(\mu)$  since, by (P2),  $\{p_\ell : \ell = 1, 2, \dots\}$  converges pointwise to  $p$  on  $D$ . Thus (1) holds with  $\rho = p'(\mu)$ .

By (P1), we have

$$\int_{[0, \infty)} e^{sx} K_\ell^\mu[dx] = e^{s\{p_\ell(\mu + \frac{s}{V_\ell}) - p_\ell(\mu)\}/(\frac{s}{V_\ell})} \quad (3.3)$$

for  $s$  in a neighbourhood of zero and hence for all  $s$  in  $[-\infty, 0]$ . Fix  $s$  and put  $\mu_\ell = \mu + \frac{s}{V_\ell}$ ; for  $\ell$  sufficiently large,  $\mu_\ell$  is in  $D$ ; moreover,  $\lim_{\ell \rightarrow \infty} \mu_\ell = \mu$ . By the convexity of  $\mu \mapsto p_\ell(\mu)$ , we have

$$(p_\ell)'_+(\mu) \leq \{p_\ell(\mu + \frac{s}{V_\ell}) - p_\ell(\mu)\}/(\frac{s}{V_\ell}) \leq (p_\ell)'_-(\mu_\ell). \quad (3.4)$$

Since  $p$  is differentiable at  $\mu$ , it follows from Griffith's Lemma that both  $\{(p_\ell)'_+(\mu)\}$  and  $\{(p_\ell)'_-(\mu_\ell)\}$  converge to  $p'(\mu)$ . Thus we have

$$\lim_{\ell \rightarrow \infty} \int_{[0, \infty)} e^{sx} K_\ell^\mu[dx] = e^{s\rho}; \quad (3.5)$$

But  $\int_{[0, \infty)} e^{sx} \delta_\rho[dx] = e^{s\rho}$ , so that (2) follows by the continuity and uniqueness theorems for the Laplace transform.

Thus we have established that if the pressure  $p$  exists and is differentiable at  $\mu$  then the sequence  $\{K_\ell^\mu : \ell = 1, 2, \dots\}$  satisfies the weak law of large numbers.

#### §4 The Heuristics of Large Deviations

Returning to the context of §2, we have the following reformulation of Theorem 1:

Suppose that the pressure  $p = \lim_{\ell \uparrow \infty} p_\ell$  exists pointwise on the interval  $D$  on which the  $p_\ell$  are defined and that  $p$  is differentiable at  $\mu$ ; then

(1) the limit  $\rho = \lim_{\ell \uparrow \infty} E_\ell^\mu[X_\ell]$  exists.

(2) Let  $g : [0, \infty) \rightarrow \mathbb{R}$  be a bounded function which is continuous at  $\rho$ ; then  $\lim_{\ell \uparrow \infty} E_\ell^\mu[g(X_\ell)] = g(\rho)$ .

Proof: (1) is a straight translation:

$$E_\ell^\mu[X_\ell] = \sum_{\omega \in \Omega} X_\ell(\omega) P_\ell^\mu[\omega] = \int_{[0, \infty)} x K_\ell^\mu[dx].$$

To prove (2), choose  $\epsilon > 0$ ; by the continuity of  $g$  at  $\rho$ , there exists a neighbourhood  $I_\rho$  of  $\rho$  on which  $|g(x) - g(\rho)| < \epsilon$ . Now

$$g(\rho) - E_\ell^\mu[g(X_\ell)] = \int_{[0, \infty)} (g(\rho) - g(x)) K_\ell^\mu[dx]. \quad (4.1)$$

Thus

$$\begin{aligned} |g(\rho) - E_\ell^\mu[g(X_\ell)]| &\leq \int_{I_\rho} |g(\rho) - g(x)| K_\ell^\mu[dx] \\ &+ \int_{I_\rho^c} |g(\rho) - g(x)| K_\ell^\mu[dx] \leq \epsilon + 2M K_\ell^\mu[I_\rho^c], \end{aligned} \quad (4.2)$$

where  $M = \sup_{x \in [0, \infty)} g(x)$ . But, by Theorem 1,  $K_\ell^\mu \rightarrow \delta_\rho$ ; this means that there exists  $\ell_0$  such that, for all  $\ell > \ell_0$ , we have  $K_\ell^\mu[I_\rho^c] < \epsilon$ . Hence, for all  $\ell > \ell_0$ ,

$$|g(\rho) - E_\ell^\mu[g(X_\ell)]| \leq \epsilon (1 + 2M); \quad (4.3)$$

but  $\epsilon$  was an arbitrary positive number, so that

$$\lim_{\ell \rightarrow \infty} E_\ell^\mu[g(X_\ell)] = g(\rho) \quad (4.4)$$

It sometimes happens in statistical mechanics that we find it interesting to introduce a perturbed grand canonical measure  $\tilde{P}_\ell^\mu$  which is conveniently defined via its expectation functional  $\tilde{E}_\ell^\mu[\cdot]$ :

$$\tilde{E}_\ell^\mu[A] = \frac{E_\ell^\mu[A e^{\beta V_\ell u(X_\ell)}]}{E_\ell^\mu[e^{\beta V_\ell u(X_\ell)}]}, \quad (4.5)$$

where  $u$  is a continuous function on  $[0, \infty)$  which is bounded above. Can we say anything about

$$\lim_{\ell \rightarrow \infty} \tilde{E}_\ell^\mu[g(X_\ell)]?$$

If the  $V_\ell$  factor were absent from the exponent, we could conclude that the limit would be the same as before:  $g(\rho)$ . However, the presence of the factor  $V_\ell$  causes the fluctuations in  $X_\ell$  to contribute to the limiting value; we expect that the answer will be  $g(\tilde{\rho})$  where  $\tilde{\rho} \neq \rho$ , in general. There are two ways of proving this; they are closely related, as we might expect.

We can introduce a perturbed Hamiltonian  $\tilde{H}_\ell(\omega) = H_\ell(\omega) + V_\ell(u \circ X_\ell)(\omega)$  and use it to define a perturbed pressure  $\tilde{p}_\ell(\mu)$ . A straightforward manipulation gives

$$\tilde{p}_\ell(\mu) = p_\ell(\mu) + \frac{1}{\beta V_\ell} \ln E_\ell^\mu[e^{\beta V_\ell u(X_\ell)}]. \quad (4.6)$$

Since

$$E_\ell^\mu[e^{\beta V_\ell u(X_\ell)}] = \int_{[0, \infty)} e^{\beta V_\ell u(x)} K_\ell^\mu[dx], \quad (4.7)$$

proof of the existence of the pressure

$$\tilde{p}(\mu) = \lim_{\ell \rightarrow \infty} \tilde{p}_\ell(\mu)$$

amounts to proving the existence of the limit

$$\lim_{\ell \rightarrow \infty} \frac{1}{\beta V_\ell} \ln \int_{[0, \infty)} e^{\beta V_\ell u(x)} K_\ell^\mu[dx];$$

conditions on  $\{K_\ell^\mu : \ell = 1, 2, \dots\}$  sufficient to ensure this were given by Varadhan [2] in a general setting, and we will give a precise statement of them in the next section. Roughly speaking, they are that there exists a function  $I^\mu(\cdot) : [0, \infty) \rightarrow [0, \infty]$  such that  $K_\ell^\mu[dx] \sim e^{-\beta V_\ell I^\mu(x)} dx$ ; in our case, the only zero of  $I^\mu$  is at  $x = \rho = p'(\mu)$ , so that  $I^\mu(\cdot)$  determines the rate at which  $P_\ell^\mu[A]$  goes to zero if  $\rho$  is not in  $A$ . Intuitively, one would expect that

$$\lim_{\ell \rightarrow \infty} \frac{1}{\beta V_\ell} \ln \int_{[0, \infty)} e^{\beta V_\ell u(x)} K_\ell^\mu[dx] = \sup_{[0, \infty)} \{u(x) - I^\mu(x)\}; \quad (4.8)$$

this is the conclusion of Varadhan's First Theorem. It follows that the perturbed pressure  $\tilde{p}(\mu)$  exists and is given by

$$\tilde{p}(\mu) = p(\mu) + \sup_{[0, \infty)} \{u(x) - I^\mu(x)\}, \quad (4.9)$$

provided that  $u$  and  $\{K_\ell^\mu\}$  satisfy the hypotheses of the theorem. If  $\tilde{p}$  is differentiable at  $\mu$ , it follows from our previous argument that

$$\lim_{\ell \rightarrow \infty} \tilde{E}_\ell^\mu[g(X_\ell)] = g(\tilde{\rho}) \quad (4.10)$$

where now  $\tilde{\rho} = \tilde{p}'(\mu)$ .

Another way of computing this limit is via Varadhan's Second Theorem: if the supremum  $\sup_{[0, \infty)} \{u(x) - I^\mu(x)\}$  is attained at an isolated point  $x^*$ , then

$$\lim_{\ell \rightarrow \infty} \tilde{E}_\ell^\mu[g(X_\ell)] = g(x^*).$$

We shall see, in our case, that the supremum is attained at an isolated point if and only if  $\tilde{p}$  is differentiable at  $\mu$  and then  $x^* = \tilde{p}'(\mu)$ .

We have seen that large deviations from the mean (deviations on the scale of  $V_\ell$ ) are of importance in the evaluation of

$$\lim_{\ell \rightarrow \infty} \frac{E_\ell^\mu[g(X_\ell)e^{\beta V_\ell u(X_\ell)}]}{E_\ell^\mu[e^{\beta V_\ell u(X_\ell)}]}.$$

It is for this reason that we are interested in the rate at which the degenerate distribution is approached; those distributions which approach the degenerate distribution exponentially fast are said to satisfy the large deviation principle. Next, we turn to the precise definition of this concept.

## §5 Varadhan's Theorems

Donsker initiated the study of singular perturbations of partial differential equations by means of functional integration; he showed how, in the case of Burger's equation, a transformation introduced by Hopf can be used to convert the equation to a linear equation which can be solved as a function space integral. The perturbation problem can then be studied by an analysis of the asymptotic behaviour of function space integrals; in the case of Burger's equation, the asymptotic analysis was carried out by Schilder [5]. Varadhan [2] showed how a class of such problems can be treated using more general families of measures on function space whose asymptotic behaviour is to be investigated; in §3 of [2], Varadhan gave an account of the asymptotic analysis in an abstract setting. Subsequently, in a sequence of papers, Donsker and Varadhan applied these methods to a wide variety of problems involving stochastic processes (a full bibliography can be found in Varadhan's monograph [6]). The method is a far-reaching generalization of

the saddle-point method or Laplace's method for one-dimensional integrals. It is our experience that whenever, in statistical mechanics, an author claims to use the saddle-point method, an efficient way of giving a rigorous proof is to check that the hypotheses of Varadhan's Theorems are verified. For that reason, we summarize here the results proved in §3 of [2].

Let  $E$  be a complete separable metric space; let  $\{K_\ell : \ell = 1, 2, \dots\}$  be a sequence of probability measures on the  $\sigma$ -field of Borel subsets of  $E$  and let  $\{V_\ell : \ell = 1, 2, \dots\}$  be a sequence of non-negative numbers such that  $V_\ell \rightarrow \infty$ . We say that  $\{K_\ell\}$  obeys the large deviation principle with constants  $\{V_\ell\}$  and rate-function  $I(\cdot)$  if there exists a function  $I : E \rightarrow [0, \infty]$  satisfying:

- (LD1):  $I(\cdot)$  is lower semi-continuous on  $E$ .  
 (LD2): For each finite  $m$ ,  $\{x : I(x) \leq m\}$  is compact.  
 (LD3): For each closed subset  $C$  of  $E$ ,  

$$\limsup_{\ell \rightarrow \infty} \frac{1}{V_\ell} \ln K_\ell[C] \leq - \inf_C I(x).$$
  
 (LD4): For each open subset  $G$  of  $E$ ,  

$$\liminf_{\ell \rightarrow \infty} \frac{1}{V_\ell} \ln K_\ell[G] \geq - \inf_G I(x).$$

For example, if  $I(\cdot)$  is a lower semi-continuous function whose level sets are compact and  $m$  is a  $\sigma$ -finite measure on  $E$  such that  $x \rightarrow e^{-I(x)}$  is integrable with respect to  $m$ , and  $\{V_\ell\}$  is a sequence of non-negative numbers such that  $V_\ell \rightarrow \infty$ , then the sequence  $\{K_\ell\}$  of probability measures defined by

$$K_\ell[A] = \frac{\int_A e^{-V_\ell I(x)} m(dx)}{\int_E e^{-\beta V_\ell I(x)} m(dx)} \quad (5.1)$$

satisfies the large deviation principle with constants  $\{V_\ell\}$  and rate-function  $I(\cdot)$ . The definition above has the advantage that it does not require the existence of a reference measure such as  $m$ . We are now in a position to state

### Varadhan's First Theorem

Let  $\{K_\ell : \ell = 1, 2, \dots\}$  be a sequence of probability measures on  $E$  obeying the large deviation principle with constants  $\{V_\ell\}$  and rate-function  $I(\cdot)$ . Then, for any continuous function  $G$  on  $E$  which is bounded above, we have

$$\lim_{\ell \rightarrow \infty} \frac{1}{V_\ell} \ln \int_{[0, \infty)} e^{V_\ell G(x)} K_\ell[dx] = \sup_E \{G(x) - I(x)\}.$$

The condition that  $G$  be bounded above can be weakened; it is enough to suppose that  $\sup\{G(x) : x \in U_{\ell \geq 1} \text{ supp } K_\ell\}$  is finite. The theorem can be extended



to cover the situation where the function  $G$  is replaced by a sequence of functions  $\{G_\ell : \ell = 1, 2, \dots\}$ ; this is Theorem 3.4 of [2].

Let  $\tilde{K}_\ell$  be defined by

$$\tilde{K}_\ell[A] = \int_A e^{V_\ell G(x)} K_\ell[dx] / \int_E e^{V_\ell G(x)} K_\ell[dx];$$

Varadhan's Second Theorem gives sufficient conditions for the sequence of perturbed measures to converge weakly to a degenerate distribution.

### Varadhan's Second Theorem

Suppose that  $\Lambda = \sup_E \{G(x) - I(x)\}$  is attained at a point  $x^*$  of  $E$  and that

$$\sup_{\{x : d(x, x^*) \geq \epsilon\}} \{G(x) - I(x)\} < \Lambda$$

for every  $\epsilon > 0$ ; if  $g$  is a bounded function on  $E$  which is continuous at  $x^*$  then

$$\lim_{\ell \rightarrow \infty} \int_E g(x) \tilde{K}_\ell[dx] = g(x^*).$$

### §6 The Upper Bound in the General Setting

In this section we return to the programme, begun in §3, of exploring the consequences of the existence of the pressure in the thermodynamic limit.

#### Theorem 2

Let  $\{K_\ell : \ell = 1, 2, \dots\}$  be a sequence of probability measures on  $[0, \infty)$  satisfying (P1) and (P2); then (LD3) holds with rate-function  $I^\mu(\cdot)$  given by

$$I^\mu(x) = p(\mu) + f(x) - \mu x,$$

where  $f(\cdot)$  is the free-energy, the Legendre transform of  $p(\cdot)$  :  $f(x) = \sup \{\mu x - p(\mu)\}$ .

Proof:

First consider an interval  $I_1 = [0, \rho_1]$  with  $\rho_1 < p'_-(\mu)$  (since  $\mu \mapsto p(\mu)$  is convex, the left-hand derivative  $p'_-(\mu)$  and the right-hand derivative  $p'_+(\mu)$  exist for all  $\mu$  in  $D$ ). For each  $\ell$  and each  $\alpha < 0$ , we have

$$\begin{aligned}
K_\ell^\mu[I_1] &= \int_{[0, \infty)} 1_{[0, \rho_1]}(x) K_\ell^\mu[dx] \leq \int_{[0, \infty)} e^{V_\ell \alpha(x - \rho_1)} K_\ell^\mu[dx] \\
&= e^{V_\ell \{p_\ell(\mu + \alpha) - p_\ell(\mu) - \alpha \rho_1\}}.
\end{aligned} \tag{6.1}$$

Thus

$$\limsup_{\ell \rightarrow \infty} \frac{1}{V_\ell} \ln K_\ell^\mu[I_1] \leq p(\mu + \alpha) - p(\mu) - \alpha \rho_1, \quad \alpha < 0. \tag{6.2}$$

It follows that

$$\begin{aligned}
\limsup_{\ell \rightarrow \infty} \frac{1}{V_\ell} \ln K_\ell^\mu[I_1] &\leq \inf_{\alpha < 0} \{p(\mu + \alpha) - p(\mu) - \alpha \rho_1\} \\
&= -\{p(\mu) + \sup_{\alpha' < \mu} \{\alpha' \rho_1 - p(\alpha')\} - \mu \rho_1\}.
\end{aligned} \tag{6.3}$$

But

$$\sup_{\alpha < \mu} \{\alpha \rho_1 - p(\alpha)\} = \sup_{\alpha} \{\alpha \rho_1 - p(\alpha)\}$$

since  $\rho < p'_-(\mu)$ ; hence

$$\limsup_{\ell \rightarrow \infty} \frac{1}{V_\ell} \ln K_\ell^\mu[I_1] \leq -I^\mu(\rho_1). \tag{6.4}$$

Next consider  $I_2 = [\rho_2, \infty)$ , where  $\rho_2 > p'_+(\mu)$ . It follows in analagous fashion that

$$\limsup_{\ell \rightarrow \infty} \frac{1}{V_\ell} \ln K_\ell^\mu[I_2] \leq -I^\mu(\rho_2). \tag{6.5}$$

Now let  $C$  be an arbitrary closed subset of  $[0, \infty)$ ; if  $C \cap [p'_-(\mu), p'_+(\mu)]$  is non-empty then  $\inf_C I^\mu(x) = 0$  and the inequality holds trivially since, for an arbitrary Borel set  $A$ , we have  $K_\ell^\mu[A] \leq 1$ ; on the other hand, if  $C \cap [p'_-(\mu), p'_+(\mu)]$  is empty, let  $(\rho_1, \rho_2)$  be the largest open interval containing  $[p'_-(\mu), p'_+(\mu)]$  which does not intersect  $C$  so that  $C \subset [0, \rho_1] \cup [\rho_2, \infty)$  and

$$\begin{aligned}
 \limsup_{\ell \rightarrow \infty} \frac{1}{V_\ell} \ln K_\ell^\mu[C] &\leq \limsup_{\ell \rightarrow \infty} \frac{1}{V_\ell} \ln \{K_\ell^\mu[I_1] + K_\ell^\mu[I_2]\} \\
 &\leq (\limsup_{\ell \rightarrow \infty} \frac{1}{V_\ell} \ln K_\ell^\mu[I_1]) \vee (\limsup_{\ell \rightarrow \infty} \frac{1}{V_\ell} \ln K_\ell^\mu[I_2]) \\
 &= (-I^\mu(\rho_1)) \vee (-I^\mu(\rho_2)) \\
 &= -\inf_C I^\mu(x),
 \end{aligned} \tag{6.6}$$

since  $x \mapsto I^\mu(x)$  is decreasing on  $[0, p'_-(\mu)]$  and increasing on  $[p'_+(\mu), \infty)$

### §7 The Berezin-Sinai Criterion for a First-Order Phase-Transition

We call an interval  $[x_1, x_2]$  of the positive real axis a first-order phase-transition segment if the free-energy function  $x \mapsto f(x)$  is linear for  $x_1 \leq x \leq x_2$ . Since  $f(x) = \sup \{\mu x - p(\mu)\}$ , each first-order phase-transition segment corresponds to a point  $\mu$  at which the grand canonical pressure is non-differentiable: there exists  $\mu$  such that

$$[x_1, x_2] = [p'_-(\mu), p'_+(\mu)].$$

On such an interval, the pressure as a function of the density is constant. In [7], Berezin and Sinai established a criterion for the existence of a first-order phase-transition segment; Dobrushin [8] simplified the proof considerably, pointing out that the criterion reduces the question of the existence of a phase-transition to the question of a "violation of the law of large numbers" in the grand canonical ensemble. Here we point out that the proof of the Berezin-Sinai criterion makes use only of the large-deviation upper bound and this, as we have seen, holds whenever the pressure exists.

### The Berezin-Sinai Criterion

Suppose that, for some  $\mu_0$  of  $\mu$ , the rate function is symmetric about some point  $x_0$ :

$$I^{\mu_0}(x_0 + y) = I^{\mu_0}(x_0 - y)$$

for all  $y$ . Suppose also that for some  $\delta > 0$ :

$$P_\ell^{\mu_0}[|X_\ell - x_0| \geq \delta] \geq c > 0$$

for all  $\ell$  sufficiently large. Then there is a first-order phase-transition at  $\mu_0$  and the interval  $[x_0 - \delta, x_0 + \delta]$  is contained in the phase-transition segment

$$[p'_-(\mu_o), p'_+(\mu_o)].$$

Proof

Let  $C = (-\infty, x_o - \delta] \cup [x_o + \delta, \infty)$ ; then, by hypothesis

$$\lim_{\sup} \frac{1}{V_\ell} \ln K_\ell^{\mu_o}[C] = 0; \quad (7.1)$$

by Theorem 2,

$$\limsup_{\ell \rightarrow \infty} \frac{1}{V_\ell} \ln K_\ell^{\mu_o}[C] \leq -\inf_C I^{\mu_o}(x) \leq 0. \quad (7.2)$$

Hence  $\inf_C I^{\mu_o}(x) = 0$ ; by the symmetry of  $I^{\mu_o}(\bullet)$  about  $x_o$ ,

$$\begin{aligned} & \inf \{I^{\mu_o}(x) : x \in [-\infty, x_o]\} \\ &= \inf \{I^{\mu_o}(x) : x \in [x_o + \delta, \infty)\} = 0 \end{aligned} \quad (7.3)$$

so that  $x_o - \delta$  and  $x_o + \delta$  must lie in  $[p'_-(\mu_o), p'_+(\mu_o)]$  and therefore

$$[x_o - \delta, x_o + \delta] \subset [p'_-(\mu_o), p'_+(\mu_o)]$$

### §8 The Lower Bound in the General Setting

We define the first-order phase-transition set  $F$  to be the union of the first-order phase-transition segments:

$$F = \cup_{\mu \in S} [p'_-(\mu), p'_+(\mu)]$$

where

$$S = \{\mu : p'_-(\mu) \neq p'_+(\mu)\}.$$

### Theorem 3

Let  $\{K_\ell : \ell = 1, 2, \dots\}$  be a sequence of probability measures on  $[0, \infty)$  satisfying (P1) and (P2); let  $G$  be an open subset of  $\text{ran } \partial p \setminus F$  where  $\partial p(\mu)$  is the sub-differential of  $p$  at  $\mu$ ; then

$$\liminf_{\ell \rightarrow \infty} \frac{1}{V_\ell} \ln K_\ell^\mu[G] \geq -\inf_G I^\mu(x).$$

Proof:

Let  $y$  be an arbitrary point of  $G$ ; choose  $\delta$  so that the neighbourhood  $B_y^\delta = (y - \delta, y + \delta)$  is contained in  $G$ ; then

$$K_\ell^\mu[G] \geq K_\ell^\mu[B_y^\delta] \int_{B_y^\delta} K_\ell^\mu[dx] = \int_{B_y^\delta} e^{-V_{\ell c_\ell^\mu}(x; \alpha)} K_\ell^{\mu+\alpha}[dx] \quad (8.1)$$

for all  $\mu$  in the domain  $D$  of  $K_\ell$  (by Lemma 2 of §3). Now choose  $\alpha$  so that  $y = p'(\mu + \alpha)$ ; this is possible because, by hypothesis,  $y$  is in  $\text{ran } \partial p \setminus F$ . Then

$$\begin{aligned} K_\ell^\mu[B_y^\delta] &= e^{-V_{\ell c_\ell^\mu}(y; \alpha)} \int_{B_y^\delta} e^{-V_{\ell c_\ell^\mu}(x-y)} K_\ell^{\mu+\alpha}[dx] \\ &\geq e^{-V_{\ell c_\ell^\mu}(y; \alpha)} e^{-V_{\ell c_\ell^\mu} \delta |\alpha|} K_\ell^{\mu+\alpha}[B_y^\delta]. \end{aligned} \quad (8.2)$$

and, by Theorem 1,  $\{K_\ell^{\mu+\alpha}\} \xrightarrow{w} \delta_y$  so that  $K_\ell^{\mu+\alpha}[B_y^\delta] > \frac{1}{2}$  for all  $\ell$  sufficiently large. Hence

$$\liminf_{\ell \rightarrow \infty} \frac{1}{V_\ell} \ln K_\ell^\mu[G] \geq -I_\ell^\mu(y) - \delta |\alpha|;$$

but  $\delta$  was an arbitrary positive number and  $y$  an arbitrary point of  $G$ ; it follows that

$$\liminf_{\ell \rightarrow \infty} \frac{1}{V_\ell} \ln K_\ell^\mu[G] \geq \sup_G (-I^\mu(y)) = -\inf_G I^\mu(y) \quad (8.3)$$

## §9 The Large Deviation Principle in the General Setting

In this section we put together the results of §6 and §8. First, we note some properties of the rate-function stemming from the convexity of the pressure  $\mu \mapsto p(\mu)$ . The free-energy  $f(\bullet)$  is the Legendre transform of  $p(\bullet)$ :

$$f(x) = \sup_{\mu} \{\mu x - p(\mu)\}. \quad (9.1)$$

Hence

$$I^\mu(x) = p(\mu) + f(x) - \mu x \geq 0.$$

We may regard  $I^\mu(\cdot)$  itself as the Legendre transform of the convex function  $\alpha \mapsto p(\mu + \alpha) - p(\mu)$ ; it follows that  $x \mapsto I^\mu(x)$  is a closed convex function and hence lower semi-continuous, so that (LD1) holds. Since  $I^\mu(x) + p(\mu + \alpha) - p(\mu) - \alpha x \geq 0$ , it follows that on the level set

$$L_m = \{x : I^\mu(x) \leq m\} \quad (9.2)$$

we have

$$\alpha x \leq m + p(\mu + \alpha) - p(\mu); \quad (9.3)$$

hence, for  $a > 0$

$$\begin{aligned} \alpha x &= \sup_{\alpha \in [-a, a]} \alpha x \leq m + \sup_{\alpha \in [-a, a]} p(\mu + \alpha) - p(\mu) \\ &< \infty. \end{aligned} \quad (9.4)$$

It follows that  $L_m$  is bounded; since  $x \mapsto I^\mu(x)$  is lower semi-continuous,  $L_m$  is closed; hence  $L_m$  is compact and so (LD2) holds.

In §6 we saw that (LD3) holds whenever the pressure exists in the thermodynamic limit; on the other hand, it is clear from §8 that more is required for (LD4) to hold since it asserts the lower bound for all open sets while the existence of the pressure suffices to establish the lower bound only for open subsets of  $\text{ran } \partial p \setminus F$ . A sufficient condition for (LD4) to hold is that  $p$  exists and is differentiable on the whole of  $R$  and that  $\text{ran } p' = [0, \infty)$ ; this is far from being necessary, however, as can be seen from the case of the free boson gas (see [3] in this volume). Nevertheless, this condition is satisfied sufficiently often to make the following theorem useful:

#### Theorem 4

Let  $\{K_\ell : \ell = 1, 2, \dots\}$  be a family of sequences of probability measures on  $[0, \infty)$  defined for all values of  $\mu$  in  $R$  and satisfying (P1) and (P2). Suppose that  $p(\cdot)$  is differentiable and that  $\text{ran } p' = [0, \infty)$ : then, for each value of  $\mu$ , the sequence  $\{K_\ell^\mu : \ell = 1, 2, \dots\}$  satisfies the large deviation principle with constants  $\{V_\ell\}$  and rate-function  $I^\mu(\cdot)$  given by  $I^\mu(x) = p(\mu) + f(x) - \mu x$  where  $f(x) = \sup_\mu \{\mu x - p(\mu)\}$ .

## §10 Remarks

I have attempted in this lecture to set out the results described in [1] in the framework established by Varadhan [2], thus showing the probabilistic consequences of the existence of the grand canonical pressure in the thermodynamic limit. Independently of [1], Ellis [9] proved a large deviation result for vector-valued random variables; his basic hypothesis is the existence of the limit of a sequence of cumulant generating functions, and Theorem 4 is a special case of his theorem.

## REFERENCES

- [1] J.T. Lewis and J.V. Pulè : The Equivalence of Ensembles in Statistical Mechanics, in Stochastic Analysis and its Applications, Proceedings, Swansea 1983 ed. A. Truman and D. Williams, LNM 1095, Springer : Heidelberg 1984.
- [2] S.R.S. Varadhan : Asymptotic Probabilities and Differential Equations, Comm. Pure Appl. Math. 19 261-286(1966).
- [3] M. van den Berg, J.T. Lewis and J.V. Pulè : Large Deviations and the Boson Gas (this volume).
- [4] L. van Hove : Quelques propriétés générales de l'intégrale de configuration d'un système de particules à une dimension, Physica 15 951-961 (1949).
- [5] M. Schilder : Some asymptotic formulae for Wiener integrals, Trans. Amer. Math. Soc. 125 63-85 (1966).
- [6] S.R.S. Varadhan : Large Deviations and Applications, Philadelphia : Society for Industrial and Applied Mathematics 1984.
- [7] F.A. Berezin and Ya. G. Sinai : Existence of a Phase-Transition of the first kind in a lattice gas with attraction between particles, Trudy Mosk. Mat. Obshch 17, 197-212 (1967).
- [8] R.L. Dobrushin : Existence of a phase-transition in models of a lattice gas, Proc. Fifth Berkeley Symposium, III, 73-87 (1967).
- [9] R.S. Ellis : Large Deviations for a general class of random vectors, Ann. Probab. 12, 1-12 (1984).

